Dynamic Link Between ECG and Clinical Data by a CORBA-Based Query Engine and Temporal Mapping

Cun Wang, M.D., M.S., Kazuhiko Ohe, M.D., Ph.D., Shigekoto Kaihara, M.D., Ph.D.
Hospital Computer Center, University of Tokyo Hospital
7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

It is important to create a dynamic link method to link distributed patient data across multiple hospitals on an "as needed" basis because the pre-defined links (an item of data has a character or group of characters that indicates the storage of another item of data) are difficult to be managed, or can only be established in part, or are not necessary to be pre-defined in many cases, especially in linking the descriptive data such as history data with the corresponding examination data across multiple hospitals. A method of linking electrocardiogram (ECG) with clinical data dynamically in a Common Object Request Broker Architecture (CORBA) environment has been achieved and verified in a real computing environment to approach to this goal. By this method, distributed patient data can be linked dynamically by a CORBA-based query engine and temporal mapping no matter where they are located on the Internet. The necessary temporal information is provided by either computing or human being. Since multiple time-axes for different databases are involved in, some temporal reasoning methods (such as mapping occurrences across temporal contexts and determining bounds for absolute occurrences, etc.) are applied to this study, and a series of temporal mappings including the first mapping, the secondary mapping, the contextual mapping, the extended mapping, the previous mapping and the next mapping are created. In comparison with the pre-defined link, the major strengths of this method are the dynamic link on an "as needed" basis, no limitation of institutional boundaries, easy creation, simplifying the data storage, and the high flexibility, etc.

INTRODUCTION

The link is a very important method for the integration of distributed data. Currently, the pre-defined link is conventionally used in both document and database management systems. In the HyperText Markup Language (HTML), links are established by indicating the symbolic address or Uniform Resource Locator (URL) of target documents. In an integrated Hospital Information System (HIS), links are usually established with a character or group of characters that indicates the storage of an item of patient data.[1] However, only such a pre-defined link is not enough in the current clinical computing because the pre-defined links are difficult to be managed, or can only be established in part, or are not necessary to be pre-defined in many cases, especially in linking clinical data including past history data, assessment data, etc. with examination data which are stored in many databases located in multiple hospitals. Along with the advance of the Internet and the increasing possibility of creating ubiquitous lifetime medical record for patients, it is important to create a link method to link distributed patient data across multiple hospitals dynamically. In this paper, we report a method and the experimental results of dynamic link between electrocardiogram (ECG) and clinical data by a query engine and temporal mapping in a Common Object Request Broker Architecture (CORBA) environment.

BACKGROUND

In our prior studies on the management of ECG data, we have created an ECG storage and retrieval system integrated with an HIS[2-3] and a Web-based interface for the object-oriented ECG database.[4] From the user interface or Internet browsers, physicians retrieve ECG data from the ECG database by indicating patient ID and the time constraints, etc. These are the independent management systems and ECG data are separated from other clinical data. In order to integrate distributed patient data together, the next important step is to link ECG with the corresponding clinical data. With the rapid advances of community health based on the Internet and the needs of hypertextual access in Electronic Medical Record (EMR), etc., we need a method to link ECG and clinical data dynamically and remotely. By using this kind of link, physicians can access corresponding ECG data from an EMR by simply clicking on clinical data that they are reading, or vice versa, no matter where data are located on the Internet or an Intranet.

In order to meet these needs, the dynamic link is necessary. Lowe et al. have created a prototype of Image Engine for retrieving and displaying clinical data such as procedure reports, laboratory data, etc. from external databases in real-time on an "as needed" basis.[5] Following this prior work, we make a research on a kind of dynamic link which focuses on hypertextual access from the contents of an EMR. By this kind of link, physicians can refer to the corresponding examination data while reading the clinical contexts no matter where data are located on the Internet - a basis of hypertextual access in the lifetime EMR. Regarding the dynamic data retrieval, the temporal query plays an important role. Many efforts have be made in the field of Artificial Intelligence in Medicine. Kahn et al. reported a context-sensitive temporal query language used to retrieve patient data from an object-oriented electronic patient medical-record system.[6] Das et al. reported a temporal query system for protocol-directed decision support.[7] Johansson et al. reported a data mapping method based on a query meta-database including temporal computations.[8] Also many researches have been made in the aspect of temporal reasoning including mapping occurrences across temporal contexts, determining bounds for absolute occurrences, deriving new occurrences from other occurrences, and deriving temporal relations between occurrences, etc.[6-11] Based on analysis of these relevant prior work on dynamic link and temporal reasoning, etc., and study on the CORBA specification, we created a query engine in a CORBA environment to match the corresponding data items dynamically between multiple databases by creating queries automatically with dealing with the uncertainty of given temporal information. If only corresponding ECG data are retrieved from the ECG database, the result is similar to the one of pre-defined link, otherwise performing further processing to make the results approximate to it.

METHODS

Capturing Temporal Information about ECG

In order to link ECG with clinical data by temporal mapping, it is necessary to capture the temporal information about ECG and keep them in the clinical database. For this study, we created a trial object-oriented clinical database which can keep temporal information about examination data (such as ECG, CT, etc.) in the corresponding clinical data items (such as past history, assessment data, and progress note, etc.). Within one hospital, the temporal information can be obtained through the integrated HIS, e.g. a time at which the ECG data are acquired (ECG-acquisition time) can be directly transferred into assessment data while ECG data are loaded in the memory. In ordering data, the temporal information about ECG is an interval from the current ECG-order entry time to the next ECG-order entry time, or $+\infty$ if there is no the next one. In some cases, the temporal information is also provided by human being, e.g. when a physician inputs a discharge summary such as "From May 1, 1996 to May 10, 1996 the patient was in hospital and had an ECG examination", the temporal information is only provided by the physician because he usually does not load ECG data again at this time. About clinical data that are related with multiple hospitals, ECG-acquisition time can also be gotten by loading ECG data remotely, but most of temporal information are provided by a patient or a physician in other hospitals, e.g. "I had an ECG examination in X hospital in 1996"; "The patient had an ECG examination in Y hospital a week ago". In order to integrate patient data efficiently, both ECG-acquisition time transferred by computing and the temporal information provided by human being are used for the dynamic link. Since temporal information is an interval or a timepoint at the different granularities, we need to use temporal reasoning methods to convert them into a pair of lower and upper bounds that cover a duration during which the corresponding ECG data are acquired. The following is a rule for the conversion: if given a timepoint, a pair of timepoints that represent the lower and upper bounds of the closed interval of uncertainty (IOU)[7] are used; if given an interval, the lower bound is equal to the lower bound of initiation, and the upper bound is equal to the upper bound of termination.

Performing Query by a CORBA-Based Engine

Since we use an open object framework to manage distributed patient data on the Web, the dynamic link mechanism is added to it. The ECG database server component and clinical database server component plug-and-play as objects in the object framework and can be interoperatable with other objects. In order to establish dynamic links between ECGs and clinical data through the framework, there should be a query engine that is divided into several parts and embedded into the client-side component, framework services and the server-side component to create a database query automatically, pass the query to the target database, execute the query in the target database and convey back the retrieved data to be displayed on the user interface, etc. Based on Object Management Group's (OMG) Interface Definition Language (IDL), we create a query pattern that comprises data name, patient ID, lower bound of the temporal information, upper bound of the temporal information, request hospital ID, and target hospital ID, etc. All these parameter values are set automatically in the client-side component according to the contexts. In order to process these parameter values especially the changeable values correctly, we handle each link point in a clinical data item, e.g. "ECG" in the item "The patient had an <u>ECG</u> examination in X hospital in 1996", as an object and other related values "1996", and "X hospital", etc. as its attribute values to be linked with it. When a physician clicks on link point "ECG", these attribute values are automatically set to the query. In order to identify a patient's data in multiple hospitals, we created a patient ID translation model to translate different patient IDs between hospitals. To unify the time systems for the temporal data on the Web, the temporal data including the lower and upper bounds of temporal information are converted into Greenwich mean time before getting into the framework, and converted to a time system similar to the local database after getting out the framework.

Since CORBA allows components and applications to communicate with one another no matter where they are located or who has designed them and extends the benefits of object-orientation across distributed heterogeneous environments including multiple languages and multiple operating systems,[12] the objects in the framework can be bound from the client-side component and above query can be passed to the server-side component remotely. If the query is successful in the target database, retrieved ECG data or clinical data are sent back to the client-side component and shown on the user interface. For the ECG data, we create their CORBA-messages by referring to A Standard Communications Protocol for Computer-Assisted Electrocardiography (SCP-ECG), our object model for the ECG database and other related materials. For the clinical data, we temporarily define their CORBA-messages according to the principle of problem-oriented medical record because currently we are using a trial clinical database and the CORBA message information reference models of Health Level Seven (HL7) are not available yet.

Matching ECG Data by Temporal Mapping

Since multiple time-axes for ECG and clinical databases are involved in, we applied some temporal reasoning methods (such as mapping occurrences across temporal contexts and determining bounds for absolute occurrences, etc.) to this study. The term "temporal mapping" in the field of Artificial Intelligence is used here to emphasize the characteristics of this method, in which occurrences are intelligently matched between multiple and remote databases.

1. The first mapping. In the object-oriented ECG database at our hospital, the ECG-acquisition time is modeled at the granularity of second (sec.),[2-3] which is the most common granularity to identify an item of data from others in the clinical computing. The times of taking ECG examination are different among the different patients. Some patients only have ECG examination one time during one year while some patients have many times within one month. If given a temporal information, all recordings of ECG data which are within the lower and upper bounds are retrieved from the ECG database. Since both a patient's frequency of ECG examination and the granularity of given temporal information affect retrieval precision rate, there are three possible results after the first mapping. The first is that only the corresponding recording of ECG data exists within the lower and upper bounds of given temporal information and the precise ECG data are retrieved from the ECG database. This is the most ideal result. In the clinical practice, almost in all cases in which a pre-defined link can be established, this result of temporal mapping can be achieved, e.g. the ECG-assessment data can be linked with the ECG recording precisely by transferring ECG-acquisition time to them while ECG data are loaded in the memory. The second is that multiple recordings of ECG data exist within the lower and upper bounds of given temporal information. The reason is that the granularity of given temporal information is not fine enough to identify the corresponding recording of ECG data from others. The coarser granularity,[9] vague specified terms (such as midday, morning)[10] and uncertain modifier (such as possibly last year),[11] etc., which are usually recorded according to the information provided by a patient or a physician, often result in this result. The third is nothing to be matched because the inaccurate temporal information was provided by a patient or a physician, and the actual ECG examination was taken out of the lower and upper bounds of given temporal information.

2. The secondary mapping. If multiple recordings of ECG data are retrieved from the ECG database after the first mapping, the secondary mapping is necessary to identify the corresponding ECG data from all retrieved data. The corresponding recording of ECG data actually exists among the limited numbers of recordings of retrieved ECG data which acquisition times are within the lower and upper bound of given temporal information, then physicians can identify it from these data easier than from all ECG data. Since ECG-acquisition times of all recordings of ECG data within the lower and upper bounds are obtained from the retrieved ECG data in the first mapping and as the lower and upper bounds to be provided for the query again, the secondary mapping has a precise result in which only one recording of ECG data is retrieved from the ECG database.

3. The contextual mapping. It is clear that the second result of the first mapping, in which the secondary mapping is necessary, is a list of ECG-acquisition times that are within the lower and upper bounds of given temporal information. If we only show this list on the user interface for the secondary mapping, the efficiency is very low because there is no clinical context to help physicians identify the corresponding one from others. To make physicians always navigate ECG within the clinical contexts, we use each ECG-acquisition time within the lower and upper bounds of given temporal information to match and display corresponding assessment and ordering data, etc. in the clinical database, and prepare dynamic links for the secondary mapping between ECGs and clinical data. By glancing over the clinical data, physicians can easily know which one is the most likely the corresponding ECG and should be first selected for the secondary mapping. Since ECG-acquisition time is kept in assessment data and can be used to match with given ECG-acquisition time, and ECG-order data can be looked back upon the nearest ECG-order entry time, the corresponding clinical data can be precisely retrieved from clinical database [exception: no corresponding clinical data available]. This mapping is also used to match corresponding clinical data from an ECG recording that has been shown on the user interface.

4. The previous and next mappings. When a physician views an ECG, he sometimes also wants to view the previous or the next one for some purposes, e.g. making a comparison on the waveforms or measurements between two or more ECGs, etc. At this time, we need the previous and next mappings. In the case of previous mapping, the lower bound is $-\infty$ or a default time, and the upper bound is the ECG-acquisition time of the current ECG data. In the case of next mapping, the lower bound is the ECG-acquisition time of the current ECG data and the upper bound is $+\infty$. After the mapping, the ECG recording which acquisition time is the closest to the one of the current ECG data is selected.

5. The extended mapping. If the result of the first mapping is nothing to be matched because of the inaccurate temporal information, it is improper to throw away the temporal information and make physicians select the corresponding one from all ECG data of a patient. Even though the temporal information is not accurate, it is near to the target time. Searching the corresponding ECG data by extending the lower and upper bounds of current temporal information is faster than searching them from all without any reference time. In the extended mapping, we extend the lower bound to $-\infty$ and the upper bound to $+\infty$ with a value and time unit (default: 1 day) until one or more recordings of ECG data are retrieved from the ECG database. There are three possibilities for the retrieving results. The first is that only corresponding recording of ECG data is retrieved. The second is that multiple recordings of ECG data including the corresponding one are retrieved. The third is that one or more recordings of ECG data but no the corresponding one are retrieved. If the result is the second one, the secondary mapping is needed, and if the result is the third one, the previous and next mappings are necessary.

RESULTS

The experiments for this method have been done with C++ programming language, MOTIF, VERSANT (an Object-Oriented Database Management System provided by Versant Object Technology Corp.) and VisiBroker (a CORBA environment provided by Visigenic Software, Inc.). A query proxy, an ECG database sever component, a trial clinical database server component, and a user interface component, etc. are created respectively and run in the several hosts which can communicate each other via the Internet. From the user interface, clinical data such as past history data, ECG-order data, and ECG-assessment data, etc. are inputted into the trial clinical database with the temporal information about ECG from the designed boxes. If ECG data are loaded in the memory, the ECG-acquisition time is automatically transferred into clinical data, otherwise the temporal information must be provided by users. When a physician clicks on clinical data that are retrieved from the clinical database and shown on the interface, the dynamic links are established between the clinical database and the ECG database automatically. Fig.1 illustrates three examples of dynamic links, in which one link is started from an item of ECG assessment data and others are started from two items of history data about ECG examinations.

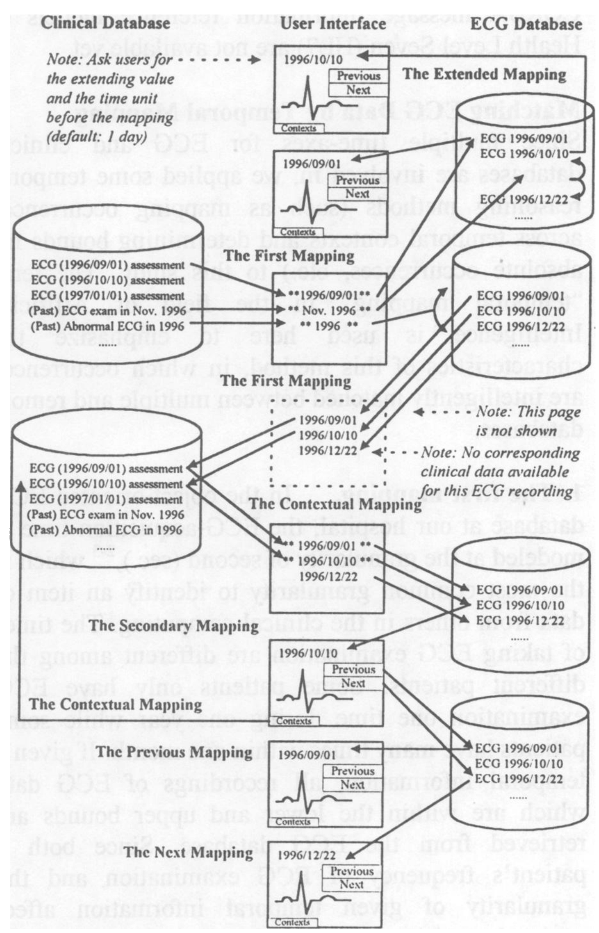

Fig.1 Temporal mappings between ECG and clinical data. Asterisks ** mean the clinical contexts.

DISCUSSION

A method of linking ECG with clinical data dynamically in a CORBA environment has been achieved and verified in a real computing environment. By this approach, distributed patient data can be linked dynamically no matter where they are located on the Internet. In this approach, the temporal mapping is a foundation, and the CORBA provides an important environment. Clinical data interchange standards, such as HL7, SCP-ECG, etc. are important references for creating CORBA-messages. About how to capture temporal and other necessary information, there are many ways including designing variable boxes for inputting specific data, and using Natural Language Processing (NLP) to extract the different information, etc. About how to keep the temporal information, etc. about objects (such as ECG, CT, etc.) in the clinical database and what is a best data model, we think that it is an important research theme needed to be studied further. Another important issue during the further studies is the evaluation of this method. For example, in using this approach in a series of situations, how often does each mapping type occur? In how many situations can mappings not be done? What are the actual cases in which extensions of the lower and upper bounds should be done in the clinical practice?

In comparison with the pre-defined link, this method has many strengths and some weaknesses. The strengths can be summarized as follows: (1)No limitation of institutional boundaries. Patient data can be linked within one or across multiple hospitals. (2)Dynamic link. The links are established in real-time on an "as needed" basis.[5] As soon as a physician clicks on the data, an item of data is linked with other one or more items of data. (3)Easy creation. The necessary temporal information for links can be provided by either computing or human being. (4)Simplifying the data storage. During data storage, what is needed to do is only keeping the necessary temporal information in clinical databases. Nothing must be done in other target databases. (5)High flexibility. At any time, patient data in any database located in any area can be linked through the object framework. The weaknesses include these: (1)Multiple steps in some cases. If the result of the first mapping is not only the corresponding data or nothing to be matched, the further mappings are necessary. (2)Imprecise links caused by inaccurate temporal information. The dynamic link can not be established very well if the inaccurate temporal information is not nearest to the acquisition time of corresponding data.

CONCLUSION

In this paper, we report a method and the experimental results of linking ECG with clinical data dynamically in a CORBA environment. We think that this method can play an important role in the researches on the EMR, especially in creating a hypermedia lifetime EMR to share and integrate patient data across multiple hospitals and clinics.

References

1. Pyle IC, Illingworth V, et al. Dictionary of computing (fourth edition). Oxford University Press, Oxford, 1996.
2. Wang C, Ohe K, Sakurai T, Nagase T, Kaihara S. Object-oriented analysis and design of an ECG storage and retrieval system integrated with an HIS. Meth Inform Med, 1996; 35(1):35-40.
3. Wang C, Ohe K, Sakurai T, Nagase T, Kaihara S. An ECG storage and retrieval system embedded in client server HIS utilizing object-oriented DB. J Med Syst, 1996; 20(1):35-43.
4. Wang C, Ohe K, Kaihara S. Applicability and problems of Java to object-oriented systems: learning from a Java programming experience of an ECG retrieval system. Japan Journal of Medical Informatics, 1997; 17(2):111-118. (in Japanese)
5. Lowe HJ, Walker WK, Vries JK. Using agent-based technology to create a cost effective, integrated, multimedia view of the electronic medical record. In: Reed M, Gardner M, ed. SCAMC 1995. Philadelphia: Harley & Belfus Inc., 1995:441-444.
6. Kahn MG, Tu S, Fagan LM. TQuery: a context-sensitive temporal query language. Comput Biomed Res, 1991; 24:401-419.
7. Das AK, Musen MA. A temporal query system for protocol-directed decision support. Meth Inform Med, 1994; 33(4):358-370.
8. Johansson B, Shahsavar N, Ahlfeldt H, Wigertz O. Database and knowledge base integration - a data mapping method for Arden syntax knowledge modules. Meth Inform Med, 1996; 35(4/5):302-308.
9. Combi C, Pinciroli F, Pozzi G. Managing different time granularities of clinical information by an interval-based temporal data model. Meth Inform Med, 1995; 34(5):458-484.
10. Keravnou ET. Temporal diagnostic reasoning based on time-objects. Artif Intell Med, 1996; 8:235-265.
11. Kahn MG, Fagan LM, Tu S. Extensions to the time-oriented database model to support temporal reasoning in medical expert systems. Meth Inform Med, 1991; 30(1):4-14.
12. Siegel J. CORBA fundamentals and programming. John Wiley & Sons, Inc., New York, 1996.